



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁷ : C12N 15/11, C07K 19/00, 16/46, C12N 1/19, 15/81	A2	(11) International Publication Number: WO 00/24884 (43) International Publication Date: 4 May 2000 (04.05.00)
(21) International Application Number: PCT/EP99/08323 (22) International Filing Date: 22 October 1999 (22.10.99) (30) Priority Data: PCT/EP98/06991 27 October 1998 (27.10.98) EP 99303118.6 22 April 1999 (22.04.99) EP (71) Applicant (for AE AU BB CA CY GB GD GH GM IE IL KE LC LK LS MN MW NZ SD SG SL SZ TT TZ UG ZA ZW only): UNILEVER PLC [GB/GB]; Unilever House, Blackfriars, London EC4P 4BQ (GB). (71) Applicant (for all designated States except AE AU BB CA CY GB GD GH GM IE IL IN KE LC LK LS MN MW NZ SD SG SL SZ TT TZ UG US ZA ZW): UNILEVER NV [NL/NL]; Weena 455, NL-3013 AL Rotterdam (NL). (71) Applicant (for IN only): HINDUSTAN LEVER LIMITED [IN/IN]; Hindustan Lever House, 165/166 Backbay Reclamation, Maharashtra, Mumbai 400 020 (IN). (72) Inventors; and (75) Inventors/Applicants (for US only): FRENKEN, Leon, Gerardus, Joseph [NL/GB]; (GB). HOWELL, Steven [GB/GB]; Unilever Research Colworth, Colworth House,		Sharnbrook, Bedford, Bedfordshire MK44 1LQ (GB). VAN DER VAART, Jan, Marcel [NL/NL]; Unilever Research Vlaardingen, Olivier van Noortlaan 120, NL-3133 AT Vlaardingen (NL). (74) Agent: EVANS, Jacqueline, Gail, Victoria; Unilever PLC, Patent Dept., Colworth House, Sharnbrook, Bedford, Bedfordshire MK44 1LQ (GB). (81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG). Published <i>Without international search report and to be republished upon receipt of that report.</i>
(54) Title: ANTIGEN-BINDING PROTEINS (57) Abstract <p>Use of a polypeptide linker group, the amino acid sequence of which group confers restricted conformational flexibility, as a linking group to link binding units, preferably antigen binding units comprising heavy chain variable domains derived from an immunoglobulin naturally devoid of light chains, in a multivalent binding protein.</p>		

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

ANTIGEN-BINDING PROTEINSFIELD OF THE INVENTION

5 The present invention relates to the preparation of multivalent and multispecific binding proteins. In particular, the invention relates to the preparation of antigen binding proteins comprising a plurality of binding units linked in series by means of intervening polypeptide linker groups, the amino acid
10 sequence of which linker groups confer restricted conformational flexibility.

BACKGROUND OF THE INVENTION

15 There is considerable interest in the preparation of multivalent and/or multispecific antigen binding proteins. Antigen binding proteins which are multivalent (that is, comprise more than one antigen binding site), more especially those which are also multispecific (where the antigen binding sites have differing
20 antigen specificities) have found particular application in the fields of diagnosis or therapy, for example, where the construction of binding proteins having binding activity against both target site and diagnostic or therapeutic agent allows for targeted delivery of the diagnostic or therapeutic agent to the
25 intended site of action. Other uses for which multivalent and multispecific binding proteins have been proposed include assays, such as immunoassays and agglutination assays, and purification processes.

30 Those multivalent, multispecific antigen binding proteins which have been described in the literature to date rely, in general, on the association of antibody light and heavy chain variable domains for the formation of the antigen binding site.

35 Thus, constructs comprising two or more polypeptide chains are described in WO 94/09131 (Scotgen Limited) and WO 97/14719

(Unilever) and WO 97/38102 (Unilever); multivalent molecules comprising two or more single chain Fv molecules linked together are described in WO 93/11161 (Enzon Inc.) and WO 94/13806 (Dow Chemical Co.).

5

WO 94/04678 (Casterman et al) describes immunoglobulins capable of exhibiting the functional properties of classical, four chain, immunoglobulins but which comprise two heavy polypeptide chains only and are naturally devoid of light polypeptide chains. Fragments corresponding to isolated VH domains or to VH dimers linked by the hinge disulphide are also disclosed. These immunoglobulins, which may be isolated from Camelids, do not rely on the association of heavy and light chain variable domains for the formation of the antigen-binding site; instead, 10 the heavy chain variable domain (hereinafter VHH) alone forms the complete antigen binding site, constituting a single domain binding site. 15

In their later patent application, WO 96/34103, Casterman et al disclose multivalent, multispecific constructs comprising VHH fragments combined with a linker sequence. Suitable linker sequences disclosed and exemplified are derived from sequences corresponding to the hinge domain of an immunoglobulin devoid of light chains. 20

25

In the Applicant's co-pending patent publication number WO 99/23221, published after the priority date of the present application, there are disclosed multivalent, multispecific antigen-binding proteins comprising a polypeptide comprising in series two or more single domain binding units which are preferably variable domains of a heavy chain derived from an immunoglobulin naturally devoid of light chains. The individual single domain binding units may suitably be linked by means of peptide linkers, preferably flexible peptide linkers, which 30 allow the variable domains to flex in relation to each other 35

with the aim of ensuring that they can bind to multiple antigenic determinants simultaneously.

There remains a continuing need for the development of improved
5 methods for producing multivalent and/or multispecific binding proteins, especially antigen binding proteins. In particular, there is commercial interest in producing molecules which not only have improved binding activity but which also can be produced economically on a large scale.

10

SUMMARY OF THE INVENTION

In a first aspect, the invention provides the use of a polypeptide group, the amino acid sequence of which group
15 confers restricted conformational flexibility, as a linking group to link binding units in a multivalent binding protein.

The invention also provides a multivalent binding protein comprising a plurality of binding units linked by means of
20 intervening polypeptide linker groups, the amino acid sequence of which linker group confers restricted conformational flexibility.

The invention further provides a nucleotide sequence encoding a
25 multivalent antigen binding protein according to the invention and cloning and expression vectors comprising such nucleotide sequences. Also provided are host cells transformed with vectors comprising such nucleotide sequences.

30 As used herein, a 'multivalent binding protein' is a protein which has more than one binding units which allow for specific binding with a molecule partner in a binding pair. Included within this are bivalent, trivalent and so on. Examples of suitable binding units include antigen binding domains of
35 antibodies, binding domains of receptors such as hormone receptors, lectins, enzymes, and cell adhesion molecules. A

'multivalent antigen binding protein' is a protein which has more than one antigen binding unit.

An 'antigen binding unit' is any structure which exhibits antigen-binding activity. This may be an antibody or an immunologically active fragment thereof. An 'antibody' refers to an immunoglobulin which may be derived from natural sources or synthetically produced. Unless indicated otherwise, 'antibody' and 'immunoglobulin' are used synonymously throughout this specification.

An antibody fragment is a portion of a whole antibody which retains the ability to exhibit antigen-binding activity. The antigen binding site may be formed through association of antibody light and heavy chain variable domains or may comprise individual antibody variable domains, constituting a single domain binding site.

Suitable fragments include Fab (comprising an antibody light chain associated with the V_H and C_{H1} domains of an antibody heavy chain), Fv (comprising the variable domains of antibody heavy and light chains associated with each other) and scFv (comprising an antibody V_H domain linked to a V_L domain by a flexible peptide linker) fragments. Where the antigen binding site comprises a single variable domain, this may be a heavy chain variable domain, most suitably a heavy chain variable domain derived from an immunoglobulin naturally devoid of light chains.

'Restricted conformational flexibility' relates to restriction of movement of the antigen binding units about the backbone of the intervening polypeptide linker group.

The present invention may be more fully understood with reference to the following description, when read together with the accompanying drawings. For convenience, an antigen binding

protein comprising two single binding units is hereinafter referred to as a 'bi-head'.

BRIEF DESCRIPTION OF THE DRAWINGS

5

Figure 1 shows a nucleotide sequence of the PstI-BstEII insert of plasmid pUR4640, encoding the heavy chain variable domain of an anti-RR6 antibody (denoted R9) from a llama.

10

Figure 2 shows the nucleotide sequence of the PstI-BstEII insert of plasmid pUR4601, encoding the heavy chain variable domain of an anti-hCG antibody (denoted H14) from a llama.

15

Figure 3 shows a map of plasmid pUR4619.

20

Figure 4 shows the nucleotide sequence within plasmid pUR4619 which encodes an anti-hCG-anti-RR6 bispecific biheaded antigen-binding protein (denoted HI4-R9), missing the first 4 and last 3 amino acids.

25

Figure 5 shows the A405 signals of an ELISA to determine bispecificity of various HI4-R9 biheads.

30

Figure 6 shows the scores achieved in a rapid assay technology (RAT) format assay following the detection of 1 IU/ml hCG (human chorionic gonadotrophin protein) with various anti-hCG-anti-RR6 bihead antigen binding proteins derived from a llama wherein the anti-hCG and anti-RR6 fragments are linked as follows (see Example 1.5):

Number	Linker	
1	no linker (directly attached)	
2	G-T-S-G-S	(SEQ. ID NO. 1)
3	S-S-S-A-S-A-S-S-A	(SEQ. ID NO. 2)
4	G-S-P-G-S-P-G	(SEQ. ID NO. 3)
5	A-T-T-T-G-S-S-P-G-P-T	(SEQ. ID NO. 4)
6	A-N-H-S-G-N-A-S	(SEQ. ID NO. 5)

Figure 7 shows a comparison of the sensitivity of detection of hCG in a RAT assay using various biheads (see Example 1.5).

5

DETAILED DESCRIPTION OF THE INVENTION

The invention is based on the unexpected finding that by using a polypeptide linking group conferring restricted conformational flexibility to link together antigen binding units, multivalent antigen binding proteins having advantageous binding affinity, as demonstrated by their increased sensitivity of diagnosis and detection, are obtained. Furthermore, constructs according to the invention may conveniently be produced at high yields economically and efficiently on a scale appropriate for industrial use.

As is apparent from the discussion of the background to the invention above, to the extent that multivalent antigen binding constructs comprising separate binding units linked together have been described at all in the literature, the linking means has been provided by flexible peptide groups. Flexibility of conformation in the linker group has been considered desirable in order to allow the multivalent construct to assume the correct orientation to allow simultaneous binding of multiple antigens.

Surprisingly, the present inventors have found that by restricting the conformational flexibility of the linking polypeptide group, multivalent antigen binding constructs having improved binding affinity may be obtained. This is entirely
5 contrary to the teaching in the art that the linking group should desirably be flexible.

The invention is applicable to the preparation of multivalent antigen binding constructs comprising antigen binding units
10 where the antigen binding site is formed through association of antibody light and heavy chain variable domains. Preferably, however, the constructs prepared according to the invention comprise a plurality of single domain binding units, more particularly a plurality of heavy chain variable domains derived
15 from an immunoglobulin naturally devoid of light chains such as may be obtained from lymphoid cells, especially peripheral blood lymphocytes, bone marrow cells or spleen cells derived from Camelids as described in WO 94/04678 (Casterman et al) discussed above. An advantage of using single domain binding units which
20 are heavy chain variable domains derived from Camelids is that they can readily and conveniently be produced economically on a large scale, for example, using a transformed lower eukaryotic host, as described in WO 94/25591 (Unilever), described above.

25 It will be appreciated that heavy chain variable domains derived from other immunoglobulins modified ('camelised') to enable them to function as single binding domains in the same way as the heavy chain variable domains derived from Camelids may also suitably be used according to the invention.

30

Bivalent forms, that is having two antigen binding sites, of the multivalent antigen binding proteins prepared according to the invention are preferred but it will be appreciated that higher multivalent forms, which are also encompassed in the present
35 invention, may find application under suitable circumstances, for example where more than two antigens are required to bind,

for example in processes for scavenging molecules from solution or processes where close proximity of molecules form the basis of an assay.

- 5 Structural features which may suitably be incorporated into the linking polypeptide group in order to achieve the effect of restricting conformational flexibility according to the purposes of the invention would readily suggest themselves to those skilled in the art.

10

Accordingly, in one embodiment, the linker group preferably comprises one or more proline residues.

- Without wishing to be bound by theory, it is generally thought
15 that the presence of a proline residue in a peptide sequence encourages the amino acid backbone of the peptide to adopt a beta-turn structural configuration, with the peptide backbone changing direction about the proline residue. Linker groups comprising other sequence features which promote the formation
20 of a beta-turn configuration in the peptide backbone, such as peptide linkers containing valine residues or constrained residues such as 8-bicyclic and 5,9-bicyclic tripeptide units (see, for example, Johannesson et al, J. Med. Chem., 42, 601-608 (1999), may also suitably find application in the present
25 invention.

- In another embodiment, peptide linker groups derived from naturally occurring proteins such as cell wall proteins (CWP), in particular, CWP1, or cellobishydrolases (CBH), such as CBH1P,
30 which serve to restrict conformational flexibility or linker groups showing at least 50% homology thereto as determined by the ALIGN program of Dayhoff et al (1983), Methods Enzymol., 91, 524-545, may also suitably be used according to the invention.

- 35 Peptide linker groups which encode a glycosylation binding site and/or are resistant to proteolytic attack may also suitably be

employed. Here, the presence of a carbohydrate attached to the amino acid residues has the effect of restricting the flexibility of the peptide backbone.

- 5 Conveniently, the polypeptide linking group according to the invention comprises from 4 to 30 amino acid residues, preferably from 5 to 15 amino acid residues.

Preferred polypeptide linking groups according to the invention
10 comprise an amino acid sequence selected from:

S-S-S-A-S-A-S-S-A, (SEQ. ID NO. 2)

G-S-P-G-S-P-G, (SEQ. ID NO. 3)

A-T-T-T-G-S-S-P-G-P-T (SEQ. ID NO. 4)

15

It will be appreciated that although the invention has been described primarily by reference to antigen binding proteins, it is equally applicable to proteins comprising other binding units as described above. References to antigen binding proteins will
20 accordingly be understood to refer also to such other proteins unless the context dictates otherwise.

Multivalent antigen binding proteins according to the invention may be prepared by transforming a host by incorporating a gene
25 encoding the polypeptide as set forth above and expressing said gene in said host.

Suitably the host or hosts may be selected from prokaryotic bacteria, such as Gram-negative bacteria, for example *E. coli*,
30 and Gram-positive bacteria, for example *B. subtilis* or lactic acid bacteria, lower eukaryotes such as yeasts, for example belonging to the genera *Saccharomyces*, *Kluyveromyces*, *Hansenula* or *Pichia*, or moulds such as those belonging to the genera *Aspergillus* or *Trichoderma*.

35

Preferred hosts for use in connection with the present invention are the lower eukaryotic moulds and yeasts.

5 Techniques for synthesising genes, incorporating them into hosts and expressing genes in hosts are well known in the art and the skilled person would readily be able to put the invention into effect using common general knowledge.

10 Methods for producing antibody fragments or functionalised fragments thereof derived from the heavy chain immunoglobulin of *Camelidae* using a transformed lower eukaryotic host are described, for example in patent application WO 94/25591 and such techniques may suitably be applied to prepare constructs according to the present invention.

15

Proteins according to the invention may be recovered and purified using conventional techniques such as affinity chromatography, ion exchange chromatography or gel filtration chromatography.

20

The activity of the multivalent binding proteins according to the invention may conveniently be measured by standard techniques known in the art such as enzyme-linked immunoadsorbant assay (ELISA), radioimmune assay (RIA) or by
25 using biosensors.

The following examples are produced by way of illustration only. Techniques used for the manipulation and analysis of nucleic acid materials were performed as described in Sambrook et al,
30 Molecular Cloning, Cold Spring Harbor Press, New York, 2nd Ed., (1989) unless otherwise indicated.

Restriction sites are underlined.

HC-V denotes heavy chain variable domain.

35

EXAMPLESExample 1 Self Assembling Llama Bi-heads Containing Linker Peptides on Latex to Assay hCG

5

1.1 Construction of Llama Bi-heads with Various Linkers

a) Induction of humeral immune responses in llama

- 10 Male llamas were immunised with a water in oil emulsion (1:9 V/V, antigen in water: Specol (Bokhout et al, Vet. Immunol. Immunopath., 2:, 491-500 (1981)) subcutaneously and intramuscularly. Per immunisation site 0.75-1.5 ml water in oil emulsion was inoculated containing 100:g antigen. The antigens
15 used were: hCG (Sigma), azo-dye RR6 (ICI) which was coupled to BSA via its reactive triazine group. Immunisations were performed according to the following time table: The second immunisation was performed three weeks after the first. The third was performed two weeks after the second immunisation.
20 The immune response was followed by antigen specific ELISAs.

- The anti-RR-6 response was measured by using Nunc Covalink plates, which were coated with the azo-dye. After incubation with (diluted) serum samples, the bound llama antibodies were
25 detected via a incubation with poly-clonal rabbit-anti-llama antiserum (obtained via immunising rabbits with llama immunoglobulins which were purified via ProtA and ProtG columns; ID-DLO), followed by an incubation with swine-anti-rabbit immunoglobulins (Dako) conjugated with alkaline phosphatase.
30 Finally the alkaline phosphatase enzyme-activity was determined after incubation with p-nitro-phenyl phosphate and the optical density was measured at 405nm. The anti-hCG response, was measured in essentially the same way using Nunc maxi-sorb plates coated with hCG.

35

b) Cloning, expressing and screening of llama HC-V fragments

i) Isolation of gene fragments encoding llama HC-V domains

5 From an immunised llama a blood sample of about 200ml was taken and an enriched lymphocyte population was obtained via Ficoll (Pharmacia) discontinuous gradient centrifugation. From these cells, total RNA was isolated by acid guanidium thiocyanate extraction (e.g. via the method described by Chomczynski and
10 Sacchi, Analytical Biochemistry, 162: 156-159 (1987). After first strand cDNA synthesis (e.g. with the Amersham first strand cDNA kit), DNA fragments encoding HC-V fragments and part of the long or short hinge region were amplified by PCR using specific primers:

15

*Pst*I

V_H - 2B 5'-AGGTSMARCTGCAGSAGTCWGG-3' (SEQ. ID NO. 6)

S = C and G, M = A and C, R = A and G, W = A and T,

20

*Hind*III

Lam-07 5'-AACAGTTAAGCTTCCGCTTGCGGCCGCGGAGCTGGGGTCTTCGCTGTGGTGCG-3'
(short hinge) (SEQ. ID NO. 7)

25

*Hind*III

Lam-08 5'-AACAGTTAAGCTTCCGCTTGCGGCCGCTGGTTGTGGTTTTGGTGTCTTGGGTT-3'
(long hinge) (SEQ. ID NO. 8)

30 Upon digestion of the PCR fragments with *Pst*I (coinciding with codon 4 and 5 of the HC-V domain, encoding the amino acids L-Q) and *Bst*EII (located at the 3'-end of the HC-V gene fragments, coinciding with the amino acid sequence Q-V-T), the DNA fragments with a length between 300 and 400bp (encoding the HC-V domain, but lacking the first three and the last three codons)
35 were purified via gel electrophoresis and isolation from the agarose gel.

ii) Construction of *Saccharomyces cerevisiae* expression plasmids encoding llama HC-V domains

Plasmids pUR4547 and pUR4548 are *Saccharomyces cerevisiae* episomal expression plasmids, derived from pSY1 (Harmsen et al., Gene, 125: 115-123, (1993)). From pSY1 the *Pst*I site, located in front of the GAL7 promoter was removed after partial digestion with *Pst*I, incubation with Klenow fragment and subsequent blunt end ligation. After transformation the desired plasmid could be selected on the basis of restriction pattern analysis. Subsequently, the *Bst*EII site in the Leu2 selection marker was removed by replacing the about 410bp *Afl*III/*Pfl*MI fragment with a corresponding fragment in which the *Bst*EII site was removed via a three step PCR mutagenesis, using the primers:

15

PCR-A:

*Pfl*MI

BOLI 1 5'-GGGAATTCCAATAGGTGCTTAGCAATCG (SEQ. ID NO. 9)

20

(*Bst*EII)

BOLI 4 5'-GACCAACGTGGTCGCCTGGCAAAACG (SEQ. ID NO. 10)

PCR-B:

(*Bst*EII)

25 BOLI 3 5'-CGTTTTGCCAGGCGACACGTTGGTC (SEQ. ID NO. 11)

*Afl*III

BOLI 2 5'-CCCCAAGCTTACATGGTCTTAAGTTGGCGT (SEQ. ID NO. 12)

30

PCR-A was performed with primers BOLI 1 and BOLI 4 and resulted in an about 130bp fragment with the *Pfl*MI restriction site at the 3'-end and the inactivated *Bst*EII site at the 5'-end. PCR-B was performed with primers BOLI 2 and BOLI 3 and resulted in an about 290bp fragment with the *Afl*III site at the 5'-end. The third PCR was with the fragments obtained from reaction A and B, together with the primers BOLI 1 and BOLI 2.

35

Finally, the about 1.8kb *SacI*-*HindIII* fragment was replaced with synthetic fragments, having sequences as presented below, resulting the plasmids pUR4547 and pUR4548, respectively.

5 - *SacI*/*HindIII* fragment of pUR4547

```

      SacI                                     (SEQ. ID NO. 13-16)
      GAGCTCATCACACAAACAAACAAACAAATGATGCTTTTGCAAGCCTTCCTT
1  -----+-----+-----+-----+-----+----- 54
10  CTCGAGTAGTGTGTTTGTGTTTGTGTTTGTGTTTACTACGAAAACGTTGGAAGGGAA
      M M L L Q A F L F
      |→      SUC2 ss

                                     PstI
15  TTCCTTTTGGCTGGTTTTCAGCCAAAATATCTGCGCAGGTGCAGCTGCAGG
55 -----+-----+-----+-----+-----+----- 105
      AAGGAAAACCGACCAAAACGTCGGTTTTATAGACGCGTCCACGTCGACGTCC
      L L A G F A A K I S A Q V Q L Q E
      |→

20                                     BstEII                                     HindIII
      AGTCATAATGAGGGACCCAGGTCACCGTCTCCTCATAATGACTTAAGCTT
106 -----+-----+-----+-----+-----+----- 155
      TCAGTATTACTCCCTGGGTCCAGTGGCAGAGGAGTATTACTGAATTTCGAA
25  E S * * G T Q V T V S S * *
      HC-V cassette                                ←|
```

and

30 - *SacI*/*HindIII* fragment of pUR4548

```

      SacI                                     (SEQ. ID NO. 17-20)
      GAGCTCATCACACAAACAAACAAACAAATGATGCTTTTGCAAGCCTTCCTT
1  -----+-----+-----+-----+-----+----- 54
35  CTCGAGTAGTGTGTTTGTGTTTGTGTTTGTGTTTACTACGAAAACGTTGGAAGGAAA
      M M L L Q A F L F
      |→      SUC2 ss

                                     PstI
40  TCCTTTTGGCTGGTTTTCAGCCAAAATATCTGCGCAGGTGCAGCTGCAGG
55 -----+-----+-----+-----+-----+----- 105
      AGGAAAACCGACCAAAACGTCGGTTTTATAGACGCGTCCACGTCGACGTCC
      L L A G F A A K I S A Q V Q L Q E
      |→

45
```


BstEII

```

AGTCATAATGAGGGACCCAGGTCACCGTCTCCTCAGAACAAAACTCATC
106-----+-----+-----+-----+-----+----- 155
TCAGTATTACTCCCTGGGTCCAGTGGCAGAGGAGTCTTGTTTTGTGAGTAG
5      S * * G T Q V T V S S E Q K L I
      HC-V cassette                ←|→      myc tail

```

HindIII

```

TCAGAAGAGGATCTGAATTAATGACTTAAGCTT
10 156-----+-----+-----+----- 188
AGTCTTCTCCTAGACTTAATTACTGAATTCGAA
      S E E D L N * *
      ←|

```

15 Both plasmids contain the *GAL7* promoter and *PGK* terminator sequences as well as the invertase (*SUC2*) signal sequence. In both plasmids the DNA sequence encoding the *SUC2* signal sequence is followed by the first 5 codons, (encoding Q-V-Q-L-Q) of the HC-V domain (including the *BstEII* site), a stuffer sequence, the

20 last six codons (encoding Q-V-T-V-S-S) of the HC-V domain. In pUR4547, this is followed by two stop codons, an *AflIII* and *HindIII* site. In pUR4548, this sequence is followed by eleven codons encoding the myc-tag, two stop codons, an *AflIII* and *HindIII* site.

25

Plasmids pUR4547 and pUR4548 were deposited under the Budapest Treaty at the Centraal Bureau voor Schimmelcultures, Baarn on 18th August 1997 with deposition numbers: CBS 100012 and CBS 100013, respectively. In accordance with Rule 28(4) EPC, or a

30 similar arrangement from a state not being a contracting state of the EPC, it is hereby requested that a sample of such deposit, when requested, will be submitted to an expert only.

Upon digesting pUR4548 with *PstI* and *BstEII*, the about 6.4kb

35 vector fragment was isolated and ligated with the *PstI*-*BstEII* fragments of about 350bp obtained as described above. After transformation of *S. cerevisiae*, via electroporation, transformants were selected from minimal medium agar plates

(comprising 0.7% yeast nitrogen base, 2% glucose and 2% agar, supplemented with the essential amino acids and bases).

iii) Screening for antigen specific HC-V domains

5

For the production of llama HC-V fragments with myc-tail, individual transformants were grown overnight in selective minimal medium (comprising 0.7% yeast nitrogen base, 2% glucose, supplemented with the essential amino acids and bases) and subsequently diluted ten times in YPGal medium (comprising 1% yeast extract, 2% bacto pepton and 5% galactose). After 24 and 48 hours of growth, the culture supernatant of the colonies was analysed by ELISA for the presence of HC-V fragments which specifically bind to the antigens hCG, RR6 in essential the same way as described above. In this case, however, the presence of specifically bound HC-V fragments was detected by incubation with monoclonal anti-myc antibodies, followed by incubation with poly-clonal rabbit-anti-mouse conjugate with alkaline phosphatase. In this way a number of anti-hCG and anti-RR6 HC-V fragments were isolated, which are:

anti-RR6:

R9 pUR4640 (see Figure 1) (SEQ. ID NO. 21-22)

25

anti-hCG (alpha unit):

H14 pUR4601 (see Figure 2) (SEQ. ID NO. 23-24)

30 c) Production of llama HC-V biheads by *S. cerevisiae*

i) Construction of episomal expression plasmids encoding anti-hCG/anti-RR6 bispecific biheads

In the anti-hCG HC-V fragment H14 (anti-alpha-subunit), the *Pst*I site was removed and a *Xho*I site was introduced via PCR, using the primers:

5 MPG158WB

*Xho*I

5'-GAATTAAGCGGCCGCCAGGTGAACTGCTCGAGTCWGGGGGA-3' (SEQ. ID NO. 25)

and

10

MPG159WB

*Bst*EII

3'-CCCTGGGTCCAGTGGCAGAGGAGTGGCAGAGGAGTCTTGTTT-5' (SEQ. ID NO. 26)

15 In this way the sequence:

*Pst*I

CAG	GTC	CAG	<u>CTG</u>	<u>CAG</u>	GAG	TCT	GGG	
Q	V	Q	L	Q	E	S	G	(SEQ. ID NO. 27)

20

became

*Xho*I

CAG	GTG	AAA	CTG	<u>CTC</u>	<u>GAG</u>	TCW	GGG	
Q	V	K	L	L	E	S	G	(SEQ. ID NO. 28)

25

Upon digesting the PCR fragments with *Xho*I and *Bst*EII, the about 330bp fragments were purified via agarose gel electrophoresis and isolation from the gel. The fragments were cloned into 30 pUR4421 (see Example 1 in WO 94/25591) which was digested with the same enzymes, resulting in pJS2 (H14). Subsequently, the about 420bp *Eag*I -*Hind*III fragment of pJS2 was isolated and ligated in the about 6.6kb *Eag*I- *Hind*III vector fragment of the pSY1 plasmid of which the *Pst*I and *Bst*EII sites were removed as 35 described above. The resulting plasmid pJS7, was digested with *Bst*EII and *Hind*III, after which the purified vector fragment was religated in the presence of a synthetic linker having the following sequence:

(SEQ. ID NO. 29-31)

```

      BstEII                      PstI                      HindIII
      <-                          MPG 160 WB (49)              ->
5  GGTCAACCGTCTCCTCACAGGTGCAGCTGCAGGAGTCACTGTAATGACTTAAGCTT
   -----+-----+-----+-----+-----+----- 55
      CCA GTGGCAGAGGAGTGTCCACGTCGACGTCCTCAGTGACATTACTGAATTGAA
      <-                          MPG 161 WB (48)              ->
      V T V S S Q V Q L Q E S L * * L K L

```

10

resulting in plasmid pJS9. Finally, the plasmid was digested with *Pst*I and *Hind*III, after which the purified vector fragments of about 7.0kb were ligated with the *Pst*I -*Hind*III fragments of about 350bp of pUR4638 and pUR4640, encoding an anti-RR6 HC-V

15 fragment denoted R9 followed by the myc-tail. The resulting *S. cerevisiae* episomal expression plasmid pUR4619 encodes a anti-hCG-anti-RR6 bispecific bihead preceded by the SUC2 signal sequence and followed by the myc-tail.

20 pUR4619: SUC2 - H14 - R9 - myc

(see Figures 3-4/SEQ. ID NO. 32-33)

Upon digesting these plasmids with *Xho*I and partially with *Bst*EII, *Xho*I-*Bst*EII fragments of about 0.7kb can be isolated and

25 subsequently cloned into the vector fragment of pUR4547 (digested with the same enzymes). In this way biheads can be obtained without the myc tail.

It will be appreciated that expression vectors can be

30 constructed in which different promoter systems, e.g. the constitutive GAPDH promoter or different signal sequences, e.g. the mating factor prepro sequence are used.

ii) Production of the HC-V biheads

35

After introducing the expression plasmid pUR4619 into *S. cerevisiae* via electroporation, transformants were selected

from minimal medium agar plates as described in part b(ii) above. For the production of biheads, the transformants were grown overnight in selective minimal medium and subsequently diluted ten times in YPGal medium. After 24 and 48 hours of growth, samples were taken for Western blot analysis. For the immuno detection of the produced biheads via Western blot analysis, monoclonal anti-myc antibodies were used, followed by incubation with poly-clonal rabbit-anti-mouse conjugate with alkaline phosphatase.

10

d) Anti-hCG/anti-RR6 bispecific biheads containing a linker peptide

i) Construction of *S. cerevisiae* episomal expression plasmids encoding anti-hCG/anti-RR6 bispecific biheads containing a linker peptide

15

Between the H14 and the R9 encoding DNA fragments synthetic linkers were introduced encoding different linker peptides. To this end the about 50 bp long *Bst*EII-*Hind*III fragment of pJS7 (see Example 1 c(i) above) was replaced by an about 50 bp long *Bst*EII-*Hind*III fragment having the following sequence:

20

MVaJA

25

*Bst*EII *Xba*I *Dra*III *Pst*I *Hind*III
5' GTCACCGTCTCTAGATGGCCACCAGGTGCAGCTGCAGGAGTCAACTTA 3'

(SEQ. ID NO. 34)

MVbJA

30

3' GCAGAGATCTACCGGTGGTCCACGTCGAGCTCCTCAGTTGAATTCCA 5'

(SEQ. ID NO. 35)

This resulted in pSJ7a. In this plasmid the about 20 bp *Pst*I-*Hind*III fragment was replaced with the about 370 bp *Pst*I-*Hind*III fragment encoding the anti-RR6 HC-V fragment R9 and/or with the myc-tail of pUR4640 (see Example 1 c(i)) and resulting in pSJ7b.

35

Upon digesting plasmid pSJ7b with *Xba*I and *Dra*III the about 7 kb vector fragment was ligated with five synthetic oligo nucleotide linker fragments presented below:

5

MV01JA 5' CTAGTGGTACTTCCGGTCCCAG 3' (SEQ. ID NO. 36)

MV02JA 3' ACCATGAAGGCCAAGG 5' (SEQ. ID NO. 37)

10

S G T S G S Q

MV03JA 5' CTAGTTCTTCATCTGCTTCTGCCTCTTCAGCCCAG 3'
(SEQ. ID NO. 38)

15 MV04JA 3' AAGAAGTAGACGAAGACGGAGAAGTCGG 5' (SEQ ID NO. 39)

S S S S A S A S S A Q

MV05JA 5' CTAGTGGTTCTCCAGGTTCAACAGGTCAG 3' (SEQ. ID NO. 40)

20

MV06JA 3' ACCAAGAGGTCCAAGTGGTCCA 5' (SEQ. ID NO. 41)

S G S P G S P G Q

25 MV07JA 5' CTAGTGCTACTACAACTGGTTCTTCACCAGGTCCAACCTCAG 3'
(SEQ. ID NO. 42)

MV08JA 3' ACGATGATGTTGACCAAGAAGTGGTCCAGGTTGA 5'
(SEQ. ID NO. 43)

30

S A T T T G S S P G P T Q

MV09JA 5' CTAGTGCTAATCATTCTGGTAATGCTTCTCAG 3'

(SEQ. ID NO. 44)

MV10JA 3' ACGATTAGTAAGACCATTACGAAGA 5' (SEQ. ID NO. 45)

5

S A N H S G N A S Q

The oligonucleotide linker fragments encode the last amino acid of the N-terminal HC-V fragment (S) and the first amino acid of the C-terminal HC-V fragment, intersected by the connecting linker peptide. This resulted in plasmids pUR5330 to 5334, respectively.

After transformation of *S. cerevisiae* with these plasmids, the production levels of the biheads were determined via Western blot analysis and a anti-hCG ELISA using anti-myc mAb for detection of the bound bihead (see Example 1 b(iii)). Production levels are presented in Table 2 below:

20 Table 2

Plasmid	Linker	Production level (mg/l)
pUR4619	None	11
pUR5330	S-G-T-S-G-S-Q	36
pUR5331	S-S-S-S-A-S-A-S-S-A-Q	49
pUR5332	S-G-S-P-G-S-P-G-Q	33
pUR5333	S-A-T-T-T-G-S-S-P-G-P-T-Q	56
pUR5334	S-A-N-H-S-G-N-A-S-Q	51

The production levels of the biheads in which the two HC-V domains are separated by a linker peptide (consisting of between 5 and 11 amino acids) were found to be 3 to 5 times higher as found for the bihead in which the two HC-V fragments are connected without a peptide linker.

Finally, the bispecificity of the biheads was demonstrated as follows:

5 PINs coated with hCG were incubated with (diluted) medium samples. Subsequently, the PINs were incubated with a RR6-alkaline phosphatase conjugate, in which the azo-dye RR6 was coupled to the alkaline phosphatase via its reactive triazine group. Finally, the alkaline phosphatase enzyme activity was
10 determined after incubation of the PINs with p-nitro-phenyl phosphate and the optical density was measured at 405nm (see Figure 5).

15 1.2 Purification of Llama Bi-heads with Various Linkers from *S. cerevisiae* Culture Media

A 5 ml column of recombinant Protein A Fast Flow Sepharose (Amersham Pharmacia Biotech) was equilibrated by washing with 10 column volumes of wash buffer (10 mM potassium phosphate, pH 6),
20 at a flow rate of 2 ml/min. The bi-head fermentation broth was loaded at 2 ml/min in an upwards direction. After loading, the column was washed with wash buffer until the OD₂₈₀ reached the baseline. Elution was carried out with a linear gradient of 0 - 40 mM citric acid pH 2.5 in the reverse direction, collecting 4
25 ml fractions into tubes containing 400 µl of a neutralising agent (1M Tris.Cl, pH 8.5) in order to minimise the effects of the acid. Peak fractions were checked for purity by running on a 12% SDS-PAGE Ready Gel (Bio-Rad) under standard denaturing conditions. Staining was with GelCode Blue (Pierce & Warriner).
30 The fractions were concentrated using Macrosep centrifugal concentrators (3 kDa molecular weight cut-off, Pall Filtron Corp.) then buffer exchanged into 10mM potassium phosphate, pH 6 using PD-10 columns (Amersham Pharmacia Biotech). The final purity of the sample was determined by carrying out a UV scan
35 from 400 - 220 nm and using the value at 280 nm to determine an

accurate concentration. The samples were then aliquoted into vials, frozen, freeze dried and stored until required.

1.3 Preparation of a Reactive Red 6-Bovine Serum Albumin
5 Conjugate

A solution of Reactive Red 6 (RR6) was made up at 10 mg/ml in phosphate buffered saline (PBS). A solution of bovine serum albumin (BSA) was made up at 10 mg/ml in PBS. 200 µl of the RR6
10 solution was added to 800 µl of the BSA solution and the resulting solution was mixed in an end over end rotary mixer for 2 hours at room temperature. RR6 that had conugated to BSA was separated from free RR6 by addition of the reaction mixture (1 ml) to a PD10 column (Pharmacia) previously washed with 10 ml of
15 PBS containing 0.1% sodium azide (PBSA). The column was then eluted by addition of PBSA (5 ml) and 1 ml aliquots were collected. The RR6-BSA conjugate eluted in fractions 4 and 5. These were pooled and the concentration of protein was determined using a BCA protein test and the concentration
20 adjusted to 2 mg/ml with PBSA.

1.4 Adsorption of Latex with Reactive Red 6-Bovine Serum
Albumin Conjugate

25 Duke blue latex was adsorbed with the RR6-BSA conjugate as follows:

To 950 µl of 10 mM borate buffer, 0.01 % merthiolate, pH 8.5 (buffer B) a 50 µl aliquot of Duke blue latex (10 % solids) was
30 added and mixed by inverting. The diluted latex was centrifuged at 8000 g for 10 minutes at room temperature, the supernatant removed and the pellet vortexed briefly. The pellet was resuspended in 900 µl of buffer B and to this 100 µl of the previously prepared RR6-BSA conjugate was added. The latex
35 solution was sonicated for 10 s using a sonic probe. The

solution containing the latex was mixed for 2 h at room temperature and then centrifuged (8000 g, 10 min at room temperature). The latex pellet was washed by resuspending in 1 ml of buffer B and centrifuged once more (8000 g, 10 min at room temperature). The pellet was then resuspended in 1 ml buffer B ready for use.

1.5 Analysis of Llama Bi-head Self Assembling on Reactive Red 6-Bovine Serum Albumin Adsorbed Latex

10

The llama bi-heads were tested by self assembling onto RR6-BSA adsorbed latex and detection of hCG in a rapid assay technology (RAT) format. This was performed by mixing the llama bi-head (5 µl of a 0.1 mg/ml solution) with RR6-BSA adsorbed latex (5 µl of 0.1 % solids) in 10 µl of PBSA to which hCG (5 µl of various concentrations) was added. The resulting solution was added to the bottom of a nitrocellulose strip (6 mm wide x 30 mm long) on which a monoclonal antibody recognising hCG had been adsorbed by plotting in a line (2.5 mg/ml) mid way up the strip. The latex-bi-head-hCG solution was allowed to flow up the nitrocellulose strip by capillary action and the strip was then washed by applying PBSA (25 µl) to the bottom of the strip. The amount of latex, captured at the plotted antibody line on the nitrocellulose strip, was quantified by measuring the absorbance through the strip.

Figure 6 shows that the llama bi-heads with linkers 3, 4 and 5 gave the highest response in RAT assays. These linkers are structurally more ordered than the comparative examples, flexible linkers 2 and 6 and result in more hCG and more latex captured in the assay. The more ordered linkers promote the correct orientation of the binding domains to achieve more optimal binding than when no linker is used. Linker 3 is derived from CWP1 and Linker 5 from CBH1P.

35

Synthetic linkers with some order (linker 4 containing 2 proline residues) can offer increased sensitivity in assays than those with little order (linker 2). Figure 7 shows that the bi-head with linker 4 can detect lower amounts (50 mIU/ml) of hCG than the bi-head with linker 2 and, hence, give a more sensitive assay for hCG.

CLAIMS

1. Use of a polypeptide group, the amino acid sequence of which group confers restricted conformational flexibility, as a linking group to link binding units in a multivalent binding protein.
2. Use according to claim 1 wherein the polypeptide linking group comprises from 4 to 30 amino acid residues.
3. Use according to claim 1 or 2 wherein the linking group comprises one or more proline residues.
4. Use according to claim 1 or 2 wherein the linking group comprises an amino acid sequence selected from:
- S-S-S-A-S-A-S-S-A,
G-S-P-G-S-P-G, or
A-T-T-T-G-S-S-P-G-P-T.
5. A multivalent binding protein comprising a plurality of binding units linked by means of intervening polypeptide linker groups, the amino acid sequence of which linker group confers restricted conformational flexibility.
6. A protein according to claim 5 wherein the binding units comprise heavy chain variable domains derived from an immunoglobulin naturally devoid of light chains.
7. A protein according to claim 5 or claim 6 wherein the antigen binding units comprise heavy chain variable domains derived from a Camelid immunoglobulin.
8. A protein according to any one of claims 5 to 7 comprising a bivalent antigen binding protein.

9. A protein according to any one of claims 5 to 8 wherein the linker group comprises from 4 to 30 amino acid residues.
10. A protein according to any one of claims 5 to 9 wherein the linker group comprises one or more proline residues.
11. A protein according to any one of claims 5 to 9 wherein the linker group comprises an amino acid sequence selected from:
- S-S-S-A-S-A-S-S-A,
G-S-P-G-S-P-G, or
A-T-T-T-G-S-S-P-G-P-T.
12. Nucleotide sequences encoding for a multivalent binding protein of any one of claims 5 to 11.
13. An expression vector comprising a nucleotide sequence according to claim 12.
14. A host cell transformed with a vector according to claim 13.

Fig.1.

```

PstI
1  CAGGTGCAGCTGCAGGAGTCAGGGGGAGGCTTGGTGCAGGCTGGGGAGTCTCTGAAACTCTCCTGTGCAGCCTCTGGAACACCTTCAGT
   +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
   90
   GTCCACGTCGACGTCCTCAGTCCCCCTCCGAACACGTCGACCCCTCAGAGACTTTGAGAGGACACGTCGGAGACCTTTGTGGAAGTCA
   +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
   Q V Q L Q E S G G G L V Q A G E S L K L S C A A S G N T F S
   [-> CDR I

KpnI
   GCGGGCTTCATGGGCTGGTACCGCCAGGCTCCAGGNAAGCAGCGGAGTTGGTCGCAACCATTAAATAGTAGGTATCACAACATATGCA
   +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
   180
   CCGCCGAAGTACCGGACCATGGCGTCCGAGGTCCTCGTGGCTCAACCCAGCGTTGGTAATTATCATCTCCATAGTGTGTTGATACGT
   +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
   G G F M G W Y R Q A P G K Q R E L V A T I N S R G I T N Y A
   [-> CDR II

EagI
   GACTTCGTGAAGGGCCGATTACCATCTCCAGAGACAATGCCAAGAGACAGTGATTTGGAAATGAACAGCCTGGAACCTGAAGACACG
   +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
   270
   CTGAAGCACTTCCCGGCTAAGTGGTAGAGGTCTCTGTACGGTCTTCTGTGCATATAAACCTTTACTTGTGCGACCTTGGACTTCTGTGC
   +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
   D F V K G R F T I S R D N A K K T V Y L E M N S L E P E D T
   [-> CDR III

BstEII
   GCCGTTTATTACTGTTACACTCACTACTTCAGATCCTACTGGGGTCAGGGGACCCAGGTCACCGTCTCCTCA
   +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
   342
   CCGCAATAATGACAATGTGAGTGATGAAGTCTAGGATGACCCCGAGTCCCGCTGGGTCCAGTGGCAGAGGAGT
   +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
   A V Y Y C Y T H Y F R S Y W G Q G T Q V T V S S
   [-> CDR III <-]

```

SUBSTITUTE SHEET (RULE 26)

2/7

Fig.2.

PstI

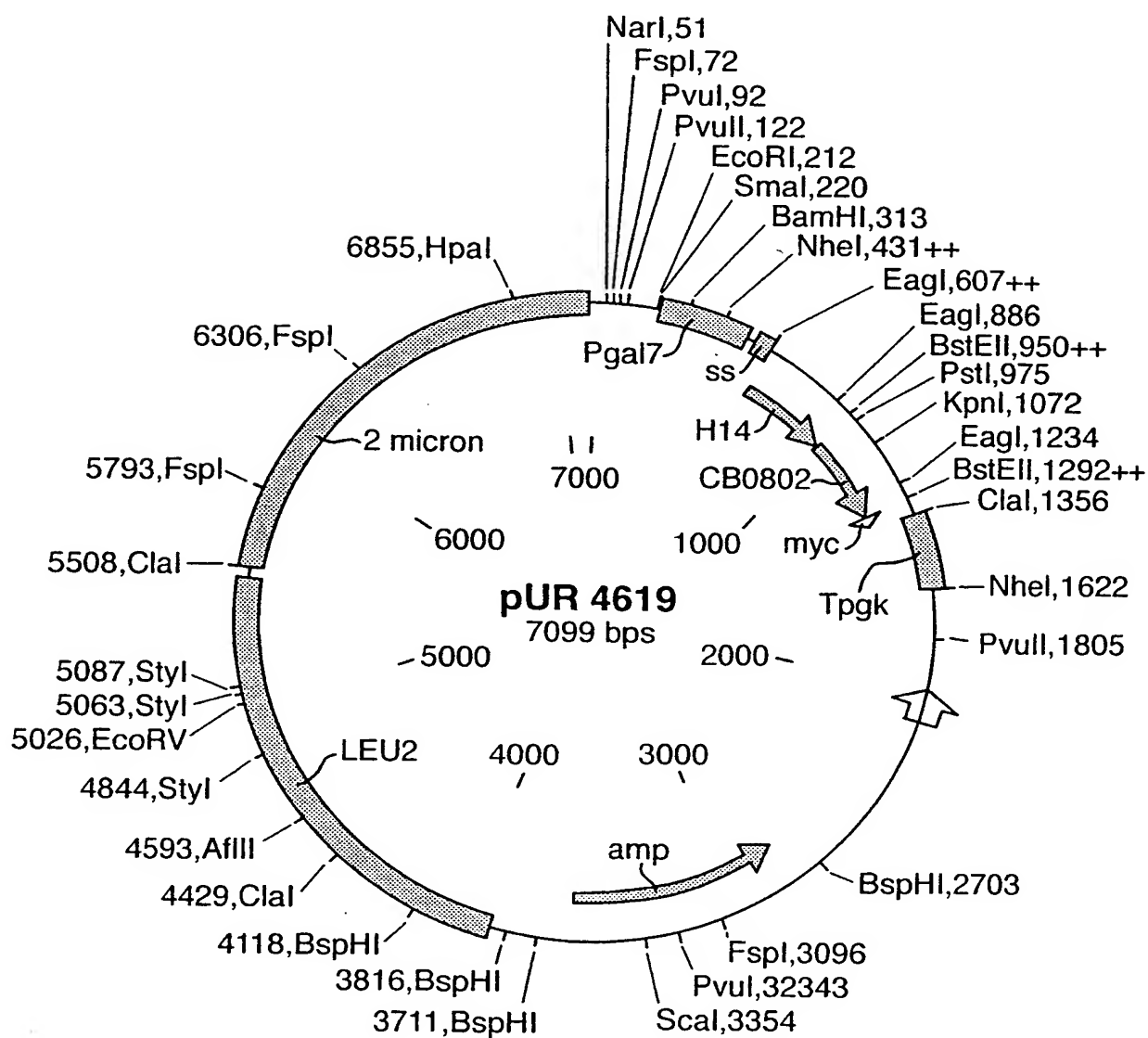
```

CAGGTGCAGCTGCAGGAGTCAGGGGAGGATTGGTGCAGCGGGGGCTCTCTGAGACTCTCCTGTGCAGCCTCTGGACGCACCGGCAGT
1  -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 90
GTCCACGTCGACGTCCTCAGTCCCCCTCCTAACACGTCGCGCCCCGAGAGACTCTGAGAGGACACGTCGGAGACCTGCGTGGCCGTC
Q V Q L Q E S G G G L V Q A G G S L R L S C A A S G R T G S
91 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 180
ACGTATGACATGGGCTGGTCCGCCAGGCTCCAGGAGGAGCGTGAGTCTGTAGCAGCTATTAACTGGGATAGTGC GCGCACATACTAT
TGCATACTGTACCCGACCAAGGGGTCCGAGGTCCCTTCTCGCACTCAGACATCGTCGATAATTGACCCCTATCACGGCGGTGTATGATA
T Y D M G W F R Q A P G K E R E S V A A I N W D S A R T Y Y
1-> CDR I <-1
181 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 270
GCAAGCTCCGTGAGGGGCGGATTCCACCATCTCCAGAGACAACGCCAAGAGACGGTGTATCTGCAATGAACAGCCTGAAACCTGAGGAC
CGTTCGAGGCACTCCCCGGCTAAGTGGTAGAGGTCTCTGTGGGTTCTTCTGCCACATAGACGTTTACTTGTGCGGACTTTGGACTCCTG
A S S V R G R F T I S R D N A K K T V Y L Q M N S L K P E D
181 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 270
<-1
BstEII
ACGGCCGTTTATACCTGTGGCGGGGGAAGGTGGTACTTGGGACTCCTGGGGCCAGGGACCCAGGTACCCGTCTCCTCA
271 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 351
TGCCGGCAAATATGGACACCGGCCCCCTTCCACCATGAACCTGAGGACCCCGGTCCCCTGGGTCCAGTGGCAGAGGAGT
T A V Y T C G A G E G T W D S W G Q G T Q V T V S S
1-> CDR III <-1

```

SUBSTITUTE SHEET (RULE 26)

Fig.3.



SUBSTITUTE SHEET (RULE 26)

Fig.4.

```

XhoI
1  CTCGAGTCAGGGGAGGATTGGTGCAGGGGGGGCTCTCTGAGACTCTCCTGTGCAGCCTCTGGACGCCGCGCAGTACGTATGACATG 90
   +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
   GAGCTCAGTCCCCCTCCTAACCAAGTCCGCCCCCGAGAGACTCTGAGAGGACACGTCGGAGACCTCGGTGGCCGTCATGCATACGTGTAC
   L E S G G G L V Q A G G S L R L S C A A S G R T G S T Y D M
                                     1-> CDR I

91  GGCTGGTCCGCCAGGCTCCAGGGAAGGAGCGTGAGTCTGTAGCAGCTATTAACTGGGATAGTGGCGGCACATACATATGAAGCTCCGTG 180
   +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
   CCGACCAAGCGGTCCGAGGTCCCTTCCCTCGCACTCAGACATCGTCGATAATTGACCCCTATCAGCGCGGTGTATGATACGTTGAGGCAC
   G W, F R Q A P G K E R E S V A A I N W D S A R T Y Y A S S V
                                     1-> CDR II
<-1

181  AGGGGCGGATTACCATCTCCAGAGACAACGCCAAGACCGGTGTATCTGCAAAATGAACAGCCTGAACCTGAGGACACGGCCGTTTAT 270
   +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
   TCCCCGGCTAAGTGGTAGAGGTCTCTGTGGGTTCTTCTGCCACATAGACGTTTACTGTGCGACTTGGACTCCTGTGCCGGCAAATA
   R G R F T I S R D N A K K T V Y L Q M N S L K P E D T A V Y
                                     <-1

271  ACCTGTGGCGGGGAAGTGGTACTTGGGACTCCTGGGGCAGGGAGCCAGGTCCCTCCTCACAGGTGCAGCTGCAGGAGTCA 360
   +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
   TGGACACCGCGCCCCCTTCCACCATGAACCTGAGGACCCCGGTCCCTGGGTCCAGTGGCAGAGGAGTGTCCACGTCGACGTCCCTCAGT
   T C G A G E G T W D S W G Q G T Q V T V S S Q V Q L Q E S
                                     1-> CDR III
                                     <-1

```

SUBSTITUTE SHEET (RULE 26)

5/7

Fig.4(Cont.)

```

361      GGGGGAGGCTTGGTGCAGGCTGGGGAGTCTCTGAAACTCTCCTGTGCAGCCTCTGGAACACACCTTCAGTGGCGGCTTCATGGGCTGGTAC
      -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
      KpnI
      CCCCCTCCGAACCAACGTCGCCACCCCTCAGAGACTTTGAGAGGACACGTCGGAGACCTTTGTGGAAGTCACCGCCGGAAGTACCCGACCATG
      G G G L V Q A G E S L K L S C A A S G N T F S G G F M G W Y
      1--> CDR I <-1
      450

451      CGCCAGGCTCCAGGGAAGCAGCGGAGTTGGTCGCAACCATTAATAGTAGAGGTATCACAACACTATGCAGACTTCGTGAAGGGCCGATTC
      -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
      540

      GCGGTCCGAGGTCCCTTCGTGCGGCTCAACCAGCGTTGGTAATTATCATCTCCATAGTGTGTTGATACGTCGTGAAGCACTTCCCGGGCTAAG
      R Q A P G K Q R E L V A T I N S R G I T N Y A D F V K G R F
      1--> CDR II <-1
      541

      ACCATCTCCAGAGACAATGCCAAGAAGACAGTGTATTGGAATGAACAGCGCTGGAACCTGAAGACACGGCGGCTTTATTACTGTTACACT
      -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
      630

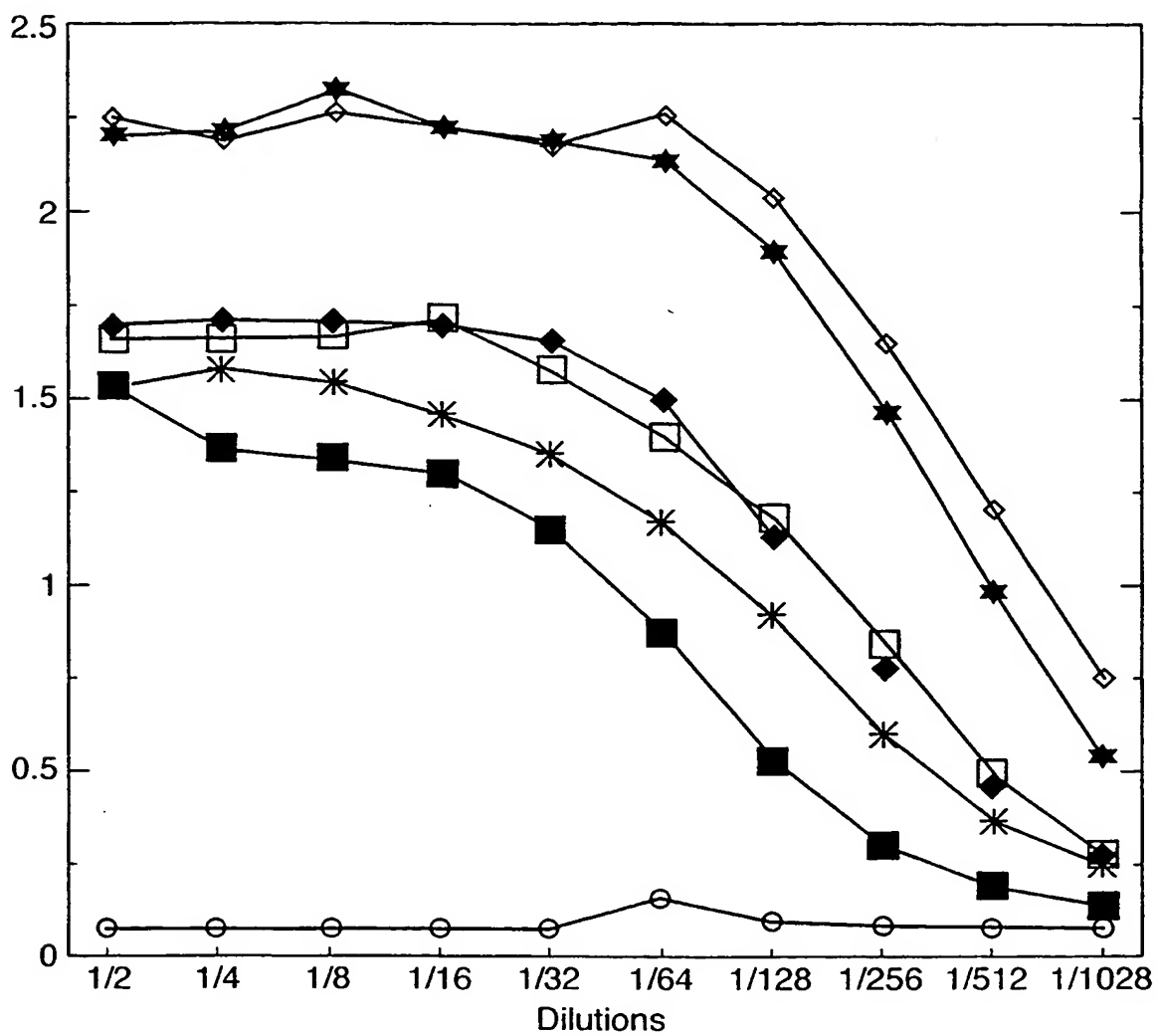
      TGGTAGAGGTCTCTGTACGGTTCTTCTGTACATAAACCTTTACTTGTGCGGACCTTGGACTTCTGTGCCGCAATAATGACAATGTGA
      T I S R D N A K K T V Y L E M N S L E P E D T A V Y Y C Y T
      1--> CDR III <-1
      631

      CACTACTTCAGATCCTACTGGGGTCAGGGGACCCAGGTCACC
      -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
      BstEII
      GTGATGAAGTCTAGGATGACCCCAAGTCCCCCTGGGTCCAGTGG
      H Y F R S Y W G Q G T Q V T
      1--> CDR III <-1
      672

```

SUBSTITUTE SHEET (RULE 26)

Fig.5.



pUR4619 ■
pUR5330 ◆
pUR5331 ★
pUR5332 □

pUR5333 ◇
pUR5334 ※
blanc ○

Fig.6.

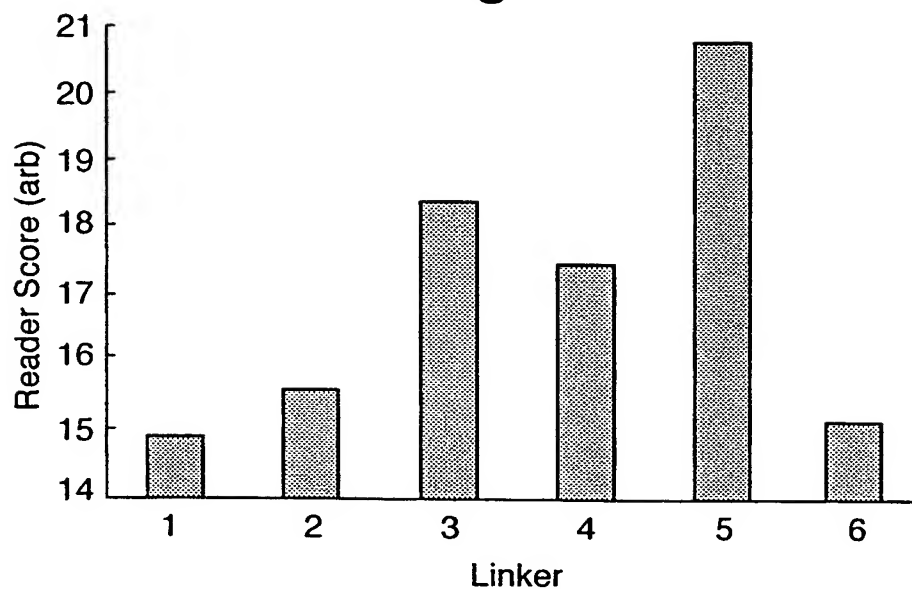
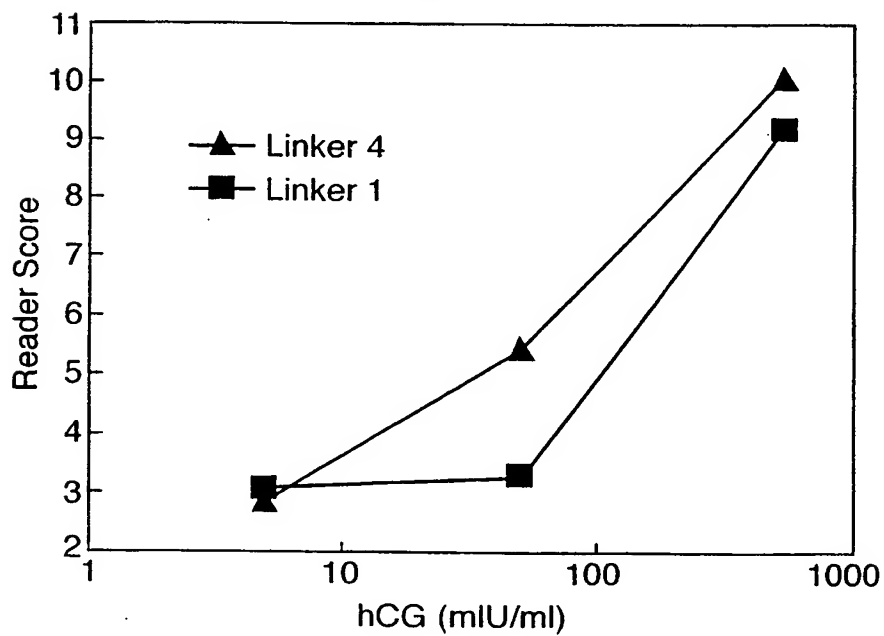


Fig.7.



SUBSTITUTE SHEET (RULE 26)

SEQUENCE LISTING

5 <110> UNILEVER PLC
UNILEVER N.V

<120> ANTIGEN BINDING PROTEINS

10 <130> T3077

<140>
<141>

15 <160> 45

<170> PatentIn Ver. 2.1

<210> 1
<211> 5

20 <212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:LINKER

25 <400> 1
Gly Thr Ser Gly Ser
1 5

30 <210> 2
<211> 9
<212> PRT
<213> Artificial Sequence

35 <220>
<223> Description of Artificial Sequence:LINKER

<400> 2

40 Ser Ser Ser Ala Ser Ala Ser Ser Ala
1 5

<210> 3
45 <211> 7
<212> PRT
<213> Artificial Sequence

<220>
50 <223> Description of Artificial Sequence:LINKER

<400> 3
Gly Ser Pro Gly Ser Pro Gly
1 5

55 <210> 4
<211> 11
<212> PRT

60 <213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:PRIMER

65 <400> 4

Ala Thr Thr Thr Gly Ser Ser Pro Gly Pro Thr
1 5 10

- 5 <210> 5
<211> 8
<212> PRT
<213> Artificial Sequence
- 10 <220>
<223> Description of Artificial Sequence:PRIMER

<400> 5
Ala Asn His Ser Gly Asn Ala Ser
15 1 5
- 20 <210> 6
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:PRIMER
25 <400> 6
aggtsmarct gcagsagtcw gg 22
- 30 <210> 7
<211> 53
<212> DNA
<213> Artificial Sequence
- 35 <220>
<223> Description of Artificial Sequence:PRIMER

<400> 7
aacagttaag cttccgcttg cggccgcgga gctggggtct tcgctgtggt gcg 53
40
- 45 <210> 8
<211> 53
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:PRIMER
50 <400> 8
aacagttaag cttccgcttg cggccgctgg ttgtgggttt ggtgtcttgg gtt 53
- 55 <210> 9
<211> 28
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:PRIMER
60 <400> 9
gggaattcca ataggtggtt agcaatcg 28
- 65

5 <210> 10
 <211> 26
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> Description of Artificial Sequence:PRIMER
 <400> 10
 10 gaccaacgtg gtcgctggc aaaacg 26
 <210> 11
 <211> 26
 15 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> Description of Artificial Sequence:PRIMER
 20 <400> 11
 cgttttgccg ggcgaccacg ttggtc 26
 25 <210> 12
 <211> 30
 <212> DNA
 <213> Artificial Sequence
 30 <220>
 <223> Description of Artificial Sequence:PRIMER
 <400> 12
 35 cccaagctt acatggtctt aagttggcgt 30
 <210> 13
 <211> 155
 <212> DNA
 40 <213> Artificial Sequence
 <220>
 <223> Description of Artificial Sequence:PLASMID
 CONSTRUCT
 45 <220>
 <221> CDS
 <222> (3)..(110)
 50 <220>
 <221> CDS
 <222> (117)..(140)
 55 <220>
 <221> CDS
 <222> (147)..(155)
 <400> 13
 60 ga gct cat cac aca aac aaa caa aac aaa atg atg ctt ttg caa gcc 47
 Ala His His Thr Asn Lys Gln Asn Lys Met Met Leu Leu Gln Ala
 1 5 10 15
 65 ttc ctt ttc ctt ttg gct ggt ttt gca gcc aaa ata tct gcg cag gtg 95
 Phe Leu Phe Leu Leu Ala Gly Phe Ala Ala Lys Ile Ser Ala Gln Val
 20 25 30

cag ctg cag gag tca taatga ggg acc cag gtc acc gtc tcc tca taatga 146
Gln Leu Gln Glu Ser Gly Thr Gln Val Thr Val Ser Ser
35 40

5 ctt aag ctt 155
Leu Lys Leu
45

10 <210> 14
<211> 36
<212> PRT
<213> Artificial Sequence

15 <223> Description of Artificial Sequence: PLASMID
CONSTRUCT

<400> 14
Ala His His Thr Asn Lys Gln Asn Lys Met Met Leu Leu Gln Ala Phe
20 1 5 10 15
Leu Phe Leu Leu Ala Gly Phe Ala Ala Lys Ile Ser Ala Gln Val Gln
20 25 30

25 Leu Gln Glu Ser
35

30 <210> 15
<211> 8
<212> PRT
<213> Artificial Sequence
<223> Description of Artificial Sequence: PLASMID
CONSTRUCT

35 <400> 15
Gly Thr Gln Val Thr Val Ser Ser
1 5

40 <210> 16
<211> 3
<212> PRT
<213> Artificial Sequence

45 <223> Description of Artificial Sequence: PLASMID
CONSTRUCT

<400> 16
Leu Lys Leu
50 1

55 <210> 17
<211> 188
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: PLASMID
CONSTRUCT

60 <220>
<221> CDS
<222> (3)..(110)

65

<220>
 <221> CDS
 <222> (117)..(173)

5 <220>
 <221> CDS
 <222> (180)..(188)

<400> 17
 10 ga gct cat cac aca aac aaa caa aac aaa atg atg ctt ttg caa gcc 47
 Ala His His Thr Asn Lys Gln Asn Lys Met Met Leu Leu Gln Ala
 1 5 10 15

15 ttc ctt ttc ctt ttg gct ggt ttt gca gcc aaa ata tct gcg cag gtg 95
 Phe Leu Phe Leu Leu Ala Gly Phe Ala Ala Lys Ile Ser Ala Gln Val
 20 25 30

20 cag ctg cag gag tca taatga ggg acc cag gtc acc gtc tcc tca gaa 143
 Gln Leu Gln Glu Ser Gly Thr Gln Val Thr Val Ser Ser Glu
 35 40 45

25 caa aaa ctc atc tca gaa gag gat ctg aat taatga ctt aag ctt 188
 Gln Lys Leu Ile Ser Glu Glu Asp Leu Xaa Xaa Lys Leu
 50 55

<210> 18
 <211> 36
 <212> PRT
 30 <213> Artificial Sequence
 <223> Description of Artificial Sequence: PLASMID
 CONSTRUCT

<400> 18
 35 Ala His His Thr Asn Lys Gln Asn Lys Met Met Leu Leu Gln Ala Phe
 1 5 10 15

40 Leu Phe Leu Leu Ala Gly Phe Ala Ala Lys Ile Ser Ala Gln Val Gln
 20 25 30

Leu Gln Glu Ser
 35

45 <210> 19
 <211> 19
 <212> PRT
 <213> Artificial Sequence
 <223> Description of Artificial Sequence: PLASMID
 50 CONSTRUCT

<400> 19
 Gly Thr Gln Val Thr Val Ser Ser Glu Gln Lys Leu Ile Ser Glu Glu
 1 5 10 15

55 Asp Leu Xaa

60 <210> 20
 <211> 3
 <212> PRT
 <213> Artificial Sequence
 <223> Description of Artificial Sequence: PLASMID
 65 CONSTRUCT

<400> 20
 Xaa Lys Leu
 1

5

<210> 21
 <211> 342
 <212> DNA
 <213> Artificial Sequence

10

<220>
 <223> Description of Artificial Sequence: PLASMID
 CONSTRUCT

15

<220>
 <221> CDS
 <222> (1)..(342)

20

<400> 21
 cag gtg cag ctg cag gag tca ggg gga ggc ttg gtg cag gct ggg gag 48
 Gln Val Gln Leu Gln Glu Ser Gly Gly Gly Leu Val Gln Ala Gly Glu
 1 5 10 15

25

tct ctg aaa ctc tcc tgt gca gcc tct gga aac acc ttc agt ggc ggc 96
 Ser Leu Lys Leu Ser Cys Ala Ala Ser Gly Asn Thr Phe Ser Gly Gly
 20 25 30

30

ttc atg ggc tgg tac cgc cag gct cca ggg aag cag cgc gag ttg gtc 144
 Phe Met Gly Trp Tyr Arg Gln Ala Pro Gly Lys Gln Arg Glu Leu Val
 35 40 45

35

gca acc att aat agt aga ggt atc aca aac tat gca gac ttc gtg aag 192
 Ala Thr Ile Asn Ser Arg Gly Ile Thr Asn Tyr Ala Asp Phe Val Lys
 50 55 60

40

ggc cga ttc acc atc tcc aga gac aat gcc aag aag aca gtg tat ttg 240
 Gly Arg Phe Thr Ile Ser Arg Asp Asn Ala Lys Lys Thr Val Tyr Leu
 65 70 75 80

45

gaa atg aac agc ctg gaa cct gaa gac acg gcc gtt tat tac tgt tac 288
 Glu Met Asn Ser Leu Glu Pro Glu Asp Thr Ala Val Tyr Tyr Cys Tyr
 85 90 95

50

act cac tac ttc aga tcc tac tgg ggt cag ggg acc cag gtc acc gtc 336
 Thr His Tyr Phe Arg Ser Tyr Trp Gly Gln Gly Thr Gln Val Thr Val
 100 105 110

tcc tca 342
 Ser Ser

55

<210> 22
 <211> 114
 <212> PRT
 <213> Artificial Sequence
 <223> Description of Artificial Sequence: PLASMID
 CONSTRUCT

60

<400> 22
 Gln Val Gln Leu Gln Glu Ser Gly Gly Gly Leu Val Gln Ala Gly Glu
 1 5 10 15

65

Ser Leu Lys Leu Ser Cys Ala Ala Ser Gly Asn Thr Phe Ser Gly Gly
 20 25 30

Phe Met Gly Trp Tyr Arg Gln Ala Pro Gly Lys Gln Arg Glu Leu Val
 35 40 45
 5 Ala Thr Ile Asn Ser Arg Gly Ile Thr Asn Tyr Ala Asp Phe Val Lys
 50 55 60
 Gly Arg Phe Thr Ile Ser Arg Asp Asn Ala Lys Lys Thr Val Tyr Leu
 65 70 75 80
 10 Glu Met Asn Ser Leu Glu Pro Glu Asp Thr Ala Val Tyr Tyr Cys Tyr
 85 90 95
 Thr His Tyr Phe Arg Ser Tyr Trp Gly Gln Gly Thr Gln Val Thr Val
 100 105 110
 15 Ser Ser

 <210> 23
 20 <211> 351
 <212> DNA
 <213> Artificial Sequence

 <220>
 25 <223> Description of Artificial Sequence: PLASMID
 CONSTRUCT

 <220>
 <221> CDS
 30 <222> (1)..(351)

 <400> 23
 cag gtg cag ctg cag gag tca ggg gga gga ttg gtg cag gcg ggg ggc 48
 Gln Val Gln Leu Gln Glu Ser Gly Gly Gly Leu Val Gln Ala Gly Gly
 35 1 5 10 15
 tct ctg aga ctc tcc tgt gca gcc tct gga cgc acc ggc agt acg tat 96
 Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Arg Thr Gly Ser Thr Tyr
 20 25 30
 40 gac atg ggc tgg ttc cgc cag gct cca ggg aag gag cgt gag tct gta 144
 Asp Met Gly Trp Phe Arg Gln Ala Pro Gly Lys Glu Arg Glu Ser Val
 35 40 45
 45 gca gct att aac tgg gat agt gcg cgc aca tac tat gca agc tcc gtg 192
 Ala Ala Ile Asn Trp Asp Ser Ala Arg Thr Tyr Tyr Ala Ser Ser Val
 50 55 60
 50 agg ggc cga ttc acc atc tcc aga gac aac gcc aag aag acg gtg tat 240
 Arg Gly Arg Phe Thr Ile Ser Arg Asp Asn Ala Lys Lys Thr Val Tyr
 65 70 75 80
 ctg caa atg aac agc ctg aaa cct gag gac acg gcc gtt tat acc tgt 288
 Leu Gln Met Asn Ser Leu Lys Pro Glu Asp Thr Ala Val Tyr Thr Cys
 85 90 95
 ggc gcg ggg gaa ggt ggt act tgg gac tcc tgg ggc cag ggg acc cag 336
 Gly Ala Gly Glu Gly Gly Thr Trp Asp Ser Trp Gly Gln Gly Thr Gln
 100 105 110
 60 gtc acc gtc tcc tca 351
 Val Thr Val Ser Ser
 115
 65

<210> 24
 <211> 117
 <212> PRT
 <213> Artificial Sequence
 5 <223> Description of Artificial Sequence: PLASMID
 CONSTRUCT

 <400> 24
 10 Gln Val Gln Leu Gln Glu Ser Gly Gly Gly Leu Val Gln Ala Gly Gly
 1 5 10 15
 Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Arg Thr Gly Ser Thr Tyr
 20 25 30
 15 Asp Met Gly Trp Phe Arg Gln Ala Pro Gly Lys Glu Arg Glu Ser Val
 35 40 45
 Ala Ala Ile Asn Trp Asp Ser Ala Arg Thr Tyr Tyr Ala Ser Ser Val
 50 55 60
 20 Arg Gly Arg Phe Thr Ile Ser Arg Asp Asn Ala Lys Lys Thr Val Tyr
 65 70 75 80
 Leu Gln Met Asn Ser Leu Lys Pro Glu Asp Thr Ala Val Tyr Thr Cys
 85 90 95
 25 Gly Ala Gly Glu Gly Gly Thr Trp Asp Ser Trp Gly Gln Gly Thr Gln
 100 105 110
 30 Val Thr Val Ser Ser
 115

 <210> 25
 35 <211> 43
 <212> DNA
 <213> Artificial Sequence

 <220>
 40 <223> Description of Artificial Sequence: PRIMER

 <400> 25
 gaattaagcg gccgcccagg tgaaactgct cgagtcwggg gga 43

 45
 <210> 26
 <211> 42
 <212> DNA
 <213> Artificial Sequence
 50
 <220>
 <223> Description of Artificial Sequence: PRIMER

 <400> 26
 55 ccctgggtcc agtggcagag gagtggcaga ggagtcttgt tt 42

 <210> 27
 <211> 24
 60 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> Description of Artificial Sequence: PRIMER
 65

<400> 27
 caggtccagc tgcaggagtc tggg 24

5 <210> 28
 <211> 24
 <212> DNA
 <213> Artificial Sequence

10 <220>
 <223> Description of Artificial Sequence:PRIMER

<400> 28
 caggtgaaac tgctcgagtc wggg 24

15

<210> 29
 <211> 55
 <212> DNA
 <213> Artificial Sequence

20

<220>
 <223> Description of Artificial Sequence:LINKER;DOUBLE
 25 STRANDED

<220>
 <221> CDS
 <222> (2)..(40)
 30

<220>
 <221> CDS
 <222> (47)..(55)

35 <400> 29
 g gtc acc gtc tcc tca cag gtg cag ctg cag gag tca ctg taatga ctt 49
 Val Thr Val Ser Ser Gln Val Gln Leu Gln Glu Ser Xaa Xaa
 1 5 10

40 aag ctt 55
 Lys Leu
 15

45 <210> 30
 <211> 13
 <212> PRT
 <213> Artificial Sequence
 <223> Description of Artificial Sequence:LINKER;DOUBLE
 50 STRANDED

<400> 30
 Val Thr Val Ser Ser Gln Val Gln Leu Gln Glu Ser Xaa
 1 5 10

55

<210> 31
 <211> 3
 <212> PRT
 <213> Artificial Sequence
 <223> Description of Artificial Sequence:LINKER;DOUBLE
 60 STRANDED

<400> 31
 65 Xaa Lys Leu

1

5 <210> 32
 <211> 672
 <212> DNA
 <213> Artificial Sequence

 10 <220>
 <223> Description of Artificial Sequence: PLASMID
 CONSTRUCT

 <220>
 <221> CDS
 15 <222> (1)..(672)

 <400> 32
 ctc gag tca ggg gga gga ttg gtg cag gcg ggg ggc tct ctg aga ctc 48
 Leu Glu Ser Gly Gly Gly Leu Val Gln Ala Gly Gly Ser Leu Arg Leu
 20 1 5 10 15
 tcc tgt gca gcc tct gga cgc acc ggc agt acg tat gac atg ggc tgg 96
 Ser Cys Ala Ala Ser Gly Arg Thr Gly Ser Thr Tyr Asp Met Gly Trp
 25 20 25 30
 ttc cgc cag gct cca ggg aag gag cgt gag tct gta gca gct att aac 144
 Phe Arg Gln Ala Pro Gly Lys Glu Arg Glu Ser Val Ala Ala Ile Asn
 35 35 40 45
 tgg gat agt gcg cgc aca tac tat gca agc tcc gtg agg ggc cga ttc 192
 Trp Asp Ser Ala Arg Thr Tyr Tyr Ala Ser Ser Val Arg Gly Arg Phe
 50 55 60
 acc atc tcc aga gac aac gcc aag aag acg gtg tat ctg caa atg aac 240
 Thr Ile Ser Arg Asp Asn Ala Lys Lys Thr Val Tyr Leu Gln Met Asn
 65 70 75 80
 agc ctg aaa cct gag gac acg gcc gtt tat acc tgt ggc gcg ggg gaa 288
 Ser Leu Lys Pro Glu Asp Thr Ala Val Tyr Thr Cys Gly Ala Gly Glu
 40 85 90 95
 ggt ggt act tgg gac tcc tgg ggc cag ggg acc cag gtc acc gtc tcc 336
 Gly Gly Thr Trp Asp Ser Trp Gly Gln Gly Thr Gln Val Thr Val Ser
 45 100 105 110
 tca cag gtg cag ctg cag gag tca ggg gga ggc ttg gtg cag gct ggg 384
 Ser Gln Val Gln Leu Gln Glu Ser Gly Gly Gly Leu Val Gln Ala Gly
 115 120 125
 gag tct ctg aaa ctc tcc tgt gca gcc tct gga aac acc ttc agt ggc 432
 Glu Ser Leu Lys Leu Ser Cys Ala Ala Ser Gly Asn Thr Phe Ser Gly
 50 130 135 140
 ggc ttc atg ggc tgg tac cgc cag gct cca ggg aag cag cgc gag ttg 480
 Gly Phe Met Gly Trp Tyr Arg Gln Ala Pro Gly Lys Gln Arg Glu Leu
 55 145 150 155 160
 gtc gca acc att aat agt aga ggt atc aca aac tat gca gac ttc gtg 528
 Val Ala Thr Ile Asn Ser Arg Gly Ile Thr Asn Tyr Ala Asp Phe Val
 60 165 170 175
 aag ggc cga ttc acc atc tcc aga gac aat gcc aag aag aca gtg tat 576
 Lys Gly Arg Phe Thr Ile Ser Arg Asp Asn Ala Lys Lys Thr Val Tyr
 180 185 190
 65

1 ttg gaa atg aac agc ctg gaa cct gaa gac acg gcc gtt tat tac tgt 624
 Leu Glu Met Asn Ser Leu Glu Pro Glu Asp Thr Ala Val Tyr Tyr Cys
 195 200 205

5 tac act cac tac ttc aga tcc tac tgg ggt cag ggg acc cag gtc acc 672
 Tyr Thr His Tyr Phe Arg Ser Tyr Trp Gly Gln Gly Thr Gln Val Thr
 210 215 220

10 <210> 33
 <211> 224
 <212> PRT
 <213> Artificial Sequence
 <223> Description of Artificial Sequence: PLASMID

15 CONSTRUCT

 <400> 33
 Leu Glu Ser Gly Gly Gly Leu Val Gln Ala Gly Gly Ser Leu Arg Leu
 1 5 10 15

20 Ser Cys Ala Ala Ser Gly Arg Thr Gly Ser Thr Tyr Asp Met Gly Trp
 20 25 30

25 Phe Arg Gln Ala Pro Gly Lys Glu Arg Glu Ser Val Ala Ala Ile Asn
 35 40 45

 Trp Asp Ser Ala Arg Thr Tyr Tyr Ala Ser Ser Val Arg Gly Arg Phe
 50 55 60

30 Thr Ile Ser Arg Asp Asn Ala Lys Lys Thr Val Tyr Leu Gln Met Asn
 65 70 75 80

 Ser Leu Lys Pro Glu Asp Thr Ala Val Tyr Thr Cys Gly Ala Gly Glu
 85 90 95

35 Gly Gly Thr Trp Asp Ser Trp Gly Gln Gly Thr Gln Val Thr Val Ser
 100 105 110

40 Ser Gln Val Gln Leu Gln Glu Ser Gly Gly Gly Leu Val Gln Ala Gly
 115 120 125

 Glu Ser Leu Lys Leu Ser Cys Ala Ala Ser Gly Asn Thr Phe Ser Gly
 130 135 140

45 Gly Phe Met Gly Trp Tyr Arg Gln Ala Pro Gly Lys Gln Arg Glu Leu
 145 150 155 160

 Val Ala Thr Ile Asn Ser Arg Gly Ile Thr Asn Tyr Ala Asp Phe Val
 165 170 175

50 Lys Gly Arg Phe Thr Ile Ser Arg Asp Asn Ala Lys Lys Thr Val Tyr
 180 185 190

55 Leu Glu Met Asn Ser Leu Glu Pro Glu Asp Thr Ala Val Tyr Tyr Cys
 195 200 205

 Tyr Thr His Tyr Phe Arg Ser Tyr Trp Gly Gln Gly Thr Gln Val Thr
 210 215 220

60 <210> 34
 <211> 48
 <212> DNA
 <213> Artificial Sequence

65

<220>
<223> Description of Artificial Sequence:SYNTHETIC
INSERT

5 <400> 34
gtcacggtct ctatgatggcc accaggtgca gctgcaggag tcaactta 48

10 <210> 35
<211> 47
<212> DNA
<213> Artificial Sequence

15 <220>
<223> Description of Artificial Sequence:SYNTHETIC
INSERT

20 <400> 35
gcagagatct accggtggtc cacgtcgagc tcctcagttg aattcga 47

25 <210> 36
<211> 23
<212> DNA
<213> Artificial Sequence

30 <220>
<223> Description of Artificial Sequence:LINKER

35 <400> 36
ctagtggtag ttccggttcc cag 23

40 <210> 37
<211> 16
<212> DNA
<213> Artificial Sequence

45 <220>
<223> Description of Artificial Sequence:LINKER

50 <400> 37
accatgaagg ccaagg 16

55 <210> 38
<211> 35
<212> DNA
<213> Artificial Sequence

60 <220>
<223> Description of Artificial Sequence:LINKER

65 <400> 38
ctagttcttc atctgcttct gcctcttcag cccag 35

60 <210> 39
<211> 28
<212> DNA
<213> Artificial Sequence

65 <220>
<223> Description of Artificial Sequence:LINKER

<400> 39
aagaagtaga cgaagacgga gaagtcgg 28

5 <210> 40
<211> 29
<212> DNA
<213> Artificial Sequence

10 <220>
<223> Description of Artificial Sequence:LINKER

<400> 40
ctagtgggtc tccaggttca ccaggtcag 29

15

<210> 41
<211> 22
<212> DNA
20 <213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:LINKER

25 <400> 41
accaagaggt ccaagtggc ca 22

<210> 42
30 <211> 41
<212> DNA
<213> Artificial Sequence

<220>
35 <223> Description of Artificial Sequence:LINKER

<400> 42
ctagtgtac tacaactggc tcttcaccag gtccaactca g 41

40

<210> 43
<211> 34
<212> DNA
<213> Artificial Sequence

45 <220>
<223> Description of Artificial Sequence:LINKER

<400> 43
50 acgatgatgt tgaccaagaa gtgggccagg ttga 34

<210> 44
<211> 32
55 <212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:LINKER

60 <400> 44
ctagtgttaa tcattctggc aatgcttctc ag 32

65 <210> 45

<211> 25
<212> DNA
<213> Artificial Sequence

5 <220>
<223> Description of Artificial Sequence:LINKER

10 <400> 45
acgattagta agaccattac gaaga

25

PCTWORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁷ : C12N 15/11, C07K 19/00, 16/46, C12N 1/19, 15/81		A3	(11) International Publication Number: WO 00/24884
			(43) International Publication Date: 4 May 2000 (04.05.00)
(21) International Application Number: PCT/EP99/08323		Sharnbrook, Bedford, Bedfordshire MK44 1LQ (GB). VAN DER VAART, Jan, Marcel [NL/NL]; Unilever Research Vlaardingen, Olivier van Noortlaan 120, NL-3133 AT Vlaardingen (NL). (74) Agent: EVANS, Jacqueline, Gail, Victoria; Unilever PLC, Patent Dept., Colworth House, Sharnbrook, Bedford, Bedfordshire MK44 1LQ (GB). (81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG). Published With international search report. (88) Date of publication of the international search report: 3 August 2000 (03.08.00)	
(22) International Filing Date: 22 October 1999 (22.10.99)			
(30) Priority Data: PCT/EP98/06991 27 October 1998 (27.10.98) EP 99303118.6 22 April 1999 (22.04.99) EP			
(71) Applicant (for AE AU BB CA CY GB GD GH GM IE IL KE LC LK LS MN MW NZ SD SG SL SZ TT TZ UG ZA ZW only): UNILEVER PLC [GB/GB]; Unilever House, Blackfriars, London EC4P 4BQ (GB).			
(71) Applicant (for all designated States except AE AU BB CA CY GB GD GH GM IE IL IN KE LC LK LS MN MW NZ SD SG SL SZ TT TZ UG US ZA ZW): UNILEVER NV [NL/NL]; Weena 455, NL-3013 AL Rotterdam (NL).			
(71) Applicant (for IN only): HINDUSTAN LEVER LIMITED [IN/IN]; Hindustan Lever House, 165/166 Backbay Reclamation, Maharashtra, Mumbai 400 020 (IN).			
(72) Inventors; and (75) Inventors/Applicants (for US only): FRENKEN, Leon, Gerardus, Joseph [NL/GB]; (GB). HOWELL, Steven [GB/GB]; Unilever Research Colworth, Colworth House,			
(54) Title: ANTIGEN-BINDING PROTEINS COMPRISING A LINKER WHICH CONFERS RESTRICTED CONFORMATIONAL FLEXIBILITY			
(57) Abstract			
Use of a polypeptide linker group, the amino acid sequence of which group confers restricted conformational flexibility, as a linking group to link binding units, preferably antigen binding units comprising heavy chain variable domains derived from an immunoglobulin naturally devoid of light chains, in a multivalent binding protein.			

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

INTERNATIONAL SEARCH REPORT

national Application No
PCT/EP 99/08323

A. CLASSIFICATION OF SUBJECT MATTER

IPC 7 C12N15/11 C07K19/00 C07K16/46 C12N1/19 C12N15/81

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 93 02198 A (TECHNICAL RESEARCH CENTRE OF FINLAND) 4 February 1993 (1993-02-04) page 11 -page 20 figures 2,4 claims	1-5,8-14
Y		6,7
Y	H. HOOGENBOOM: "Mix and match: Building manifold binding sites." NATURE BIOTECHNOLOGY, vol. 15, no. 2, February 1997 (1997-02), pages 125-126, XP002110046 New York, NY, USA page 126, right-hand column, line 30 - line 39 figure 1	6,7

	--- -/--	

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

* Special categories of cited documents:

- "A" document defining the general state of the art which is not considered to be of particular relevance
 "E" earlier document but published on or after the international filing date
 "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
 "O" document referring to an oral disclosure, use, exhibition or other means
 "P" document published prior to the international filing date but later than the priority date claimed

- "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
 "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
 "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
 "&" document member of the same patent family

Date of the actual completion of the international search

17 May 2000

Date of mailing of the international search report

24/05/2000

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Authorized officer

Nooij, F

INTERNATIONAL SEARCH REPORT

International Application No

PCT/EP 99/08323

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	M. SRISODSUK ET AL.: "Role of the interdomain linker peptide of <i>Trichoderma reesei</i> cellobiohydrolase I in its interaction with crystalline cellulose." THE JOURNAL OF BIOLOGICAL CHEMISTRY, vol. 268, no. 28, 5 October 1993 (1993-10-05), pages 20756-20761, XP002112074 Baltimore, MD, USA abstract figure 1	1-5, 9-14
X	P. HOLLIGER ET AL.: "Specific killing of lymphoma cells by cytotoxic T-cells mediated by a bispecific diabody." PROTEIN ENGINEERING, vol. 9, no. 3, March 1996 (1996-03), pages 299-305, XP002112075 Oxford, GB abstract figure 1	1,2,5,8, 9,12-14
X	A. PLÜCKTHUN ET AL.: "New protein engineering approaches to multivalent and bispecific antibody fragments." IMMUNOTECHNOLOGY, vol. 3, no. 2, June 1997 (1997-06), pages 83-105, XP004088466 Amsterdam, NL page 98, right-hand column, line 27 - line 46 figure 6E	1,2,5,8, 9
X	P. PORTOGHESE ET AL.: "Synthesis and opioid antagonist potencies of naltrexamine bivalent ligands with conformationally restricted spacers." JOURNAL OF MEDICINAL CHEMISTRY, vol. 29, no. 9, September 1986 (1986-09), pages 1650-1653, XP002112076 Washington, DC, USA abstract figure 2	1,2,5,9
A	K. ALFTHAN ET AL.: "Properties of a single-chain antibody containing different linker peptides." PROTEIN ENGINEERING, vol. 8, no. 7, July 1995 (1995-07), pages 725-731, XP002112077 Oxford, GB abstract figure 1A	1-5,9-11
	----- -/--	

INTERNATIONAL SEARCH REPORT

International Application No
PCT/EP 99/08323

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 94 13804 A (CAMBRIDGE ANTIBODY TECHNOLOGY LTD. & MEDICAL RESEARCH COUNCIL) 23 June 1994 (1994-06-23) examples claims -----	1-14
A	WO 97 14719 A (UNILEVER NV) 24 April 1997 (1997-04-24) examples claims -----	1-14

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/EP 99/08323

Patent document cited in search report		Publication date	Patent family member(s)		Publication date
WO 9302198	A	04-02-1993	FI	913434 A	17-01-1992
WO 9413804	A	23-06-1994	AU	690528 B	30-04-1998
			AU	5654894 A	04-07-1994
			CA	2150262 A	23-06-1994
			EP	0672142 A	20-09-1995
			JP	8504100 T	07-05-1996
			US	5837242 A	17-11-1998
			US	6010884 A	04-01-2000
			AT	173740 T	15-12-1998
			AU	680685 B	07-08-1997
			AU	7621494 A	10-04-1995
			CA	2169620 A	30-03-1995
			DE	69414870 D	07-01-1999
			DE	69414870 T	12-05-1999
			EP	0720624 A	10-07-1996
			ES	2126145 T	16-03-1999
			WO	9508577 A	30-03-1995
			JP	9503759 T	15-04-1997
WO 9714719	A	24-04-1997	AU	6873396 A	07-05-1997
			BR	9606706 A	06-04-1999
			CN	1173878 A	18-02-1998
			EP	0799244 A	08-10-1997
			US	5989830 A	23-11-1999

Form PCT/ISA/210 (patent family annex) (July 1992)